多数決規則の粒子系と確率モデル

伊藤　栄明

1. 現象と確率モデル

確率論は現実の現象と深くかかわりあっている．水に浮かべた花粉のこきざみな不規則運動はBrown運動として知られている．Einstein（1905,「アインシュタイン選集1」，共立出版）は思考実験からBrown運動を考え，その確率モデルを考えた．その研究は実験にもとづいたPerrin（1908）によるAvogadro数の計算にもちいられ当時問題になっていた原子論への有力な根拠をあたえた．硬貨を投げておもてが出れば正の方向に1歩，裏が出れば負の方向に1歩進むランダムウォークを連続化した確率モデルがBrown運動であると考えることができる．Brown運動の確率モデルと同等な確率モデルはEinstein以前にBachelier（1900）により株価の変動の理解のためにもちいられていた．Bachelierの論文のタイトルは「投機の理論」である．経済的にゆたかになりたいという欲求は人類社会でのふるくからのものであり，Bachelierは株価の変動の研究から基本的な確率モデルを発見した．

近代的な商品先物取引は1730年に大阪において世界で最初に始められたそうである．19世紀末，樋口一葉（1872-1896）は父親が亡くなった後の一家の生活をにない，米相場にも関心をもった（「樋口一葉日記」1894年2月23日）．Bachelier（1870-1946）は1889年に両親を亡くし，家業を引き継いだことから（1989-1991），金融市場についての知識を得たといわれている[1]．Bachelierの研究以前に金融市場というものの存在があった．経済活動以外にも選挙，言語，婚姻等人間の活動についての多くの魅力的な確率論の問題がある．人間社会の変化，新しい制度，科学技術の変化から生まれる具体的な課題は今後も新しい確率モデルをうみだしてゆくと思われる．BachelierやEinsteinほどでない研究者でも運よく面白い課題にであえば新しい確率論の分野をひらけるかもしれないと思う．

Brown運動ほど基本的でなくても，確率論には様々なモデルがある．離散的な構造のうえでの粒子の確率的相互作用について，興味を持ってきたが[5][6][8]，ここでは粒子の3体，4体の相互作用を考え，物理学における強磁性，集団遺伝学における超優性，を多数決規則の粒子系として単純化する[7]．自然現象だけでなく，人間にかかわる現象についてもこの多数決規則の粒子系の確率モデルを適用する．世界の言語の語順規則の確率的変化[9]，株価の変動[15][18]等の現象について，筆者と共同研究者による研究を中心に述べてみたい．

2. Isingモデルと超優性モデル

集団遺伝学における超優性モデルはふるくからいわれている雑種強勢をモデル化したものである[11]．例えば4体の粒子の相互作用にもとづいたモデルを考えると，集団遺伝学における超優性モデルに対応するものになる．$A_1, A_2, \cdots, A_m$なるm個の異なった遺伝子からなる系を考える．2倍体生物でN個体からなるものとする．$2N$個の遺伝子のなかからまず2個の遺伝子がランダムに選ばれ結合しA_iA_jとなるとする．のこりの$2N$-2個の遺伝子のなかから2個の遺伝子がランダムに選ばれ結合しA_kA_lとなるとする．A_iA_jとA_kA_lが衝突し，確率$1/2+s_{ij,kl}$で2個のA_iA_j，確率$1/2+s_{kl,ij}$で2個のA_kA_lとなるものとする．ここで$s_{ij,kl}=-s_{kl,ij}$とする．

$$s_{ij,kl}=\begin{cases} s/2 & \text{if } i\neq j \text{ and } k=l \\ -s/2 & \text{if } i=j \text{ and } k\neq l \\ 0 & \text{if } i\neq j \text{ and } k\neq l \\ 0 & \text{if } i=j \text{ and } k=l, \end{cases}$$

このような衝突が繰り返されてゆくとする．時間$[t, t+dt]$にこの4体の衝突がCdtの確率でおきるとし，それぞれの遺伝子の数は時刻tで$\vec{N}=(N_1, N_2,$

いとう　よしあき
情報・システム研究機構 統計数理研究所，総合研究大学院大学
〒106-8569 港区南麻布4-6-7

$\cdots, N_m)$ であるとする．粒子の総数を n とすると遺伝子頻度 N_α, $\alpha=1, 2, \cdots, m$ の平均および共分散は $\alpha, \beta=1, 2, \cdots, m$ として

$$E(\Delta N_\alpha)=2C_s\Delta t\cdot \frac{2N_\alpha}{n}\left(\frac{\sum_k N_k(N_k-1)-n(N_\alpha-1)}{(n-1)(n-2)}\right) \quad (1)$$

$$E(\Delta N_\alpha \Delta N_\beta)=\frac{4}{n(n-1)}C\Delta t N_\alpha(n\delta_{\alpha\beta}X_\beta). \quad (2)$$

となる．世代の交代は一度に行われ次世代の遺伝子頻度は現世代の遺伝子の集団からのランダムサンプリングより得られるという Fisher-Wright モデルにもとづき，フィットネスという量をもちいることにより超優性モデル（overdominance model）は通常定式化されている．粒子の相互作用という考え方によりモデル化したものがこのモデルであると考えられる．$s>0$ が小である場合は集団遺伝学において議論されている方程式と同じ拡散近似をもつと考えてよく，文献[10]による研究の結果とほぼ同じ結果を与え，超優性モデルと同様に集団に多様な遺伝子を保持する現象を説明する自然なモデルであると考えられる．このモデルでは $s>0$ である場合は小数派が有利であるが，$s<0$ である場合は多数派が有利である．

$m=2$ として A_1 から A_2，A_2 から A_1 について等しい突然変異率 r を仮定すると，$\alpha=1, 2$ について

$$E(\Delta N_\alpha)=\left(r\frac{-N_\alpha+N_{\alpha+1}}{n}+2Cs\cdot \frac{2N_\alpha}{n}\left(\frac{\sum_k N_k(N_k-1)-n(N_\alpha-1)}{(n-1)(n-2)}\right)\right)\Delta t. \quad (3)$$

となり，Fokker-Planck 方程式をもちいて系の挙動を解析することができる．$m>2$ の場合は拡散近似をもちいる方法は解析的にも数値的な解をうることも困難であり，モデルを直接シミュレーションするか，あるいは Brown 運動に基づく確率微分方程式を差分化しシミュレーションを行う．

$m=2$，$s=-1/2$ の場合はランダムに選ばれた4粒子において少数派が多数派に変化するというルールである．臨界温度から温度が下がったときに，例えば周囲に上向きスピンが多ければ下向きスピンは上向きスピンに変化してゆき磁区が形成されてゆくと考えられる．上向きスピンの割合を Brown 運動を用いた確率微分方程式により近似し解の挙動が議論されている[3][13]．相互作用によりスピンの向きが周囲の多数の方に変化してゆくものとする上記 $s<0$ の多数決によるモデルは，温度が下がってゆく過程のモデル化と考えることできる．上記 $0<s$ によるモデルは温度があがってゆき，次第に上向きスピンと下向きスピンが混在した状態に配列されてゆく過程のモデル化である．

3. 語順規則の Ising モデル

語順規則は J. Greenberg により法則のかたちにまとめられ，言語類型論として知られている．角田(1991)[16]は世界の130の言語について次に示す19個の項目についての語順規則を表にまとめた．1. S, O と V，2. 名詞と側置詞，3. 所有格と名詞，4. 指示詞と名詞，5. 数詞と名詞，6. 形容詞と名詞，7. 関係節と名詞，8. 固有名詞と普通名詞，9. 比較の表現，10. 本動詞と助動詞，11. 副詞と動詞，12. 副詞と形容詞，13. 疑問の印，14. 一般疑問文での S，V 倒置，15. 疑問詞，16. 特殊疑問分での S，V 倒置，17. 否定の印，18. 条件節と主節，19. 目的節と主節，例えば，SOV の語順をもつ言語の多くは後置詞をもち，後置詞をもつならば所有格は名詞の前にくる．日本語の助詞は後置詞ということができる．SVO の語順をもつ言語の多くは前置詞をもち，前置詞をもつ言語の所有格は名詞の後に位置することが多い．前者は日本語に近い語順をもつ言語の特徴であり，後者は日本語と逆の語順を示すものである．韓国語，ベンガル語が前者に属し，英語，タイ語は後者にあたる．階層クラスタ分析を用いると世界の言語は語順によって概ね2つのタイプに分類でき，この分類は，「2. 名詞と側置詞」の順序によって説明される[17]．

アジア・ヨーロッパ地域の言語の角田による語順の表を数値化し，最長距離法による階層クラスタ分析を適用した結果が図1である．これらは，ヨーロッパ諸言語について自然な結果を与えていると思われる．一方，ヒンヅー語，ベンガル語，パンジャブ語等の北インドの言語が，ヨーロッパ諸言語から離れて日本語の近くに分類されることは興味深い．

語順規則の変化を時間的，地理的な確率過程としてとらえることは自然なことであると思われる．確率モデルとしてとらえるという視点にたてば，言語が分裂，統合，相互作用を繰り返しながら変化してゆき，たまたま語順が非常に近い対が現れたという解釈もありうる．例えば，角田による語順表においてタミル語と日本語の間の違いは，イタリア語とスペイン語ほどに近い．このことは偶然の結果であるのか，それとも別の観点からの議論から必然性が示せるのか，興味のあるところである．日本語，タミール語，朝鮮語，蒙古語の相互の近さについて比較言語学からの研究も知られ

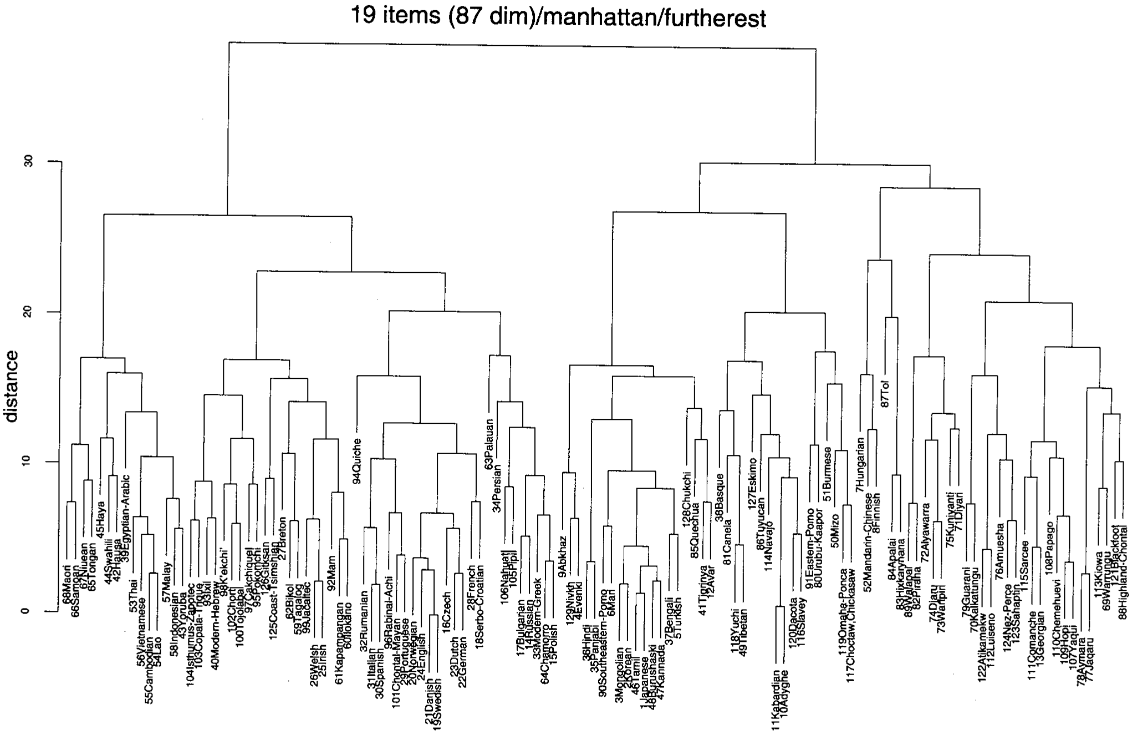

図1　語順表の階層クラスタ分析[17]

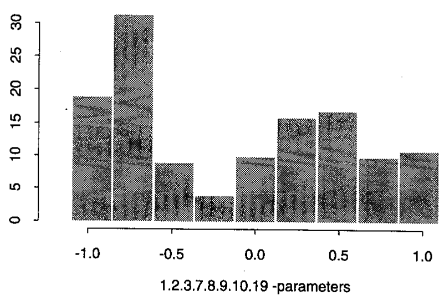

図2　8項目の算術平均の度数分布[9]

ている[4]．語順規則について集団遺伝学の場合のように中立説を仮定すれば，語順規則から言語の系統を考えることができると思われるが，各言語は地球の表面に分布しており，相互に影響し合っている．生物種とちがい系統というとらえ方だけではとらえきれないことはあきらかであろう．DNA 列を解析し生物種の系統を考える際に，変化の速い部分とおそい部分を区別して考えないと混乱した結果がえられる．同様な理由から，言語の場合語順規則と基礎語彙の両方を同時に考えるということでなく，語順規則のみにより，階層クラスタリングを考えるということにより立場がより明確になるように思われる．名詞と側置詞の項目およびそれと相関の高い項目，全部で8個の上記の項目，1，2，3，7，8，9，10，19を考え，日本語とおなじ場合を+1，逆の場合を−1とする．現実には+1，−1といいきれない中間的なものもあるが，それらは，+1と−1の間の値に数値化し，各言語について8個の項目についてのそれらの数値の算術平均を考える．図2はそれについての度数分布をあたえたものである．

これを説明する確率モデルとして3粒子の多数決によるモデルを考える[9]．系にプラスとマイナス2種類の粒子がそれぞれ n_1，n_2 個の粒子があったとする．総粒子数を n 個とし，ランダムに選ばれた3粒子のうちの2個がプラスと1個がマイナスであったとすればマイナスの1粒子がプラスに変化するものとする．同様に3粒子のうちの2個がマイナスと1個がプラスであったとすれば指定されたプラスの1粒子がマイナスに変化するものとする．3粒子とも同じものの組みあわせである場合変化はおこらないとする．この操作を繰り返して行くとき，それぞれの箱の粒子数はどのように変動するのであろうか．あきらかに最終的にはどちらかの箱に全部の粒子が集まるはずである．さらに各粒子はプラスからマイナスへある確率で突然変異するとし，マイナスからプラスへもその確率で突然変異すると仮定することにより図2において観察されたような2個の山のある度数分布がえられるが，このような確率モデルにより語順規則の変化を説明できるのではないかと思う．日本語と同じ語順を+1，逆の語順を−1とあらわしたが，各項目の変化は独立に起きるのではなく，関連する項目のプラス，マイナスと関連した形で変化するはずであり $n=8$ である多数決モ

デルを考えるのは自然であるように思う．さきにのべた，SOV の語順をもつ言語の多くは後置詞をもち，後置詞をもつならば所有格は名詞の前にくる．SVO の語順をもつ言語の多くは前置詞をもち，前置詞をもつ言語の所有格は名詞の後に位置することが多い，等のことからこのことは納得できるであろう．このモデルは語順規則の Ising モデルとでもいってよいものである．このモデルによりデータを理解できるということは，上記 8 項目の全部がプラスあるいは全部がマイナスという状態が安定であるということをある意味で示すものである．階層クラスタリングにより生成される図 1 の tree 上の 130 の言語のうえのランダムウオークを考える．8 項目の値の算術平均の値の変化は多数決によるモデルと類似の動きをする．-1 から $+1$ までの各値について訪問回数の度数分布は図 1 に類似し 2 個の山を持つものが得られる．

4. 多数派志向トレーダーによる株価の振動

「日本経済新聞」(2004 年 1 月 13 日けいざい心理学 3) に次のような記事があった．「大手銀行の為替ディーラー富樫直人さん（仮名）がもっとも頼りにしているのは同僚ディーラーがどのように動いたのかを知る行内ファイル．方向性が自分の相場観とちがってもみんなと同じ方をむいたほうが安心できる．多くの銀行はエクセルファイルで各ディーラーの動きを閲覧できるようにしている」．

トレーダーには多数派志向があると考え，多数決による粒子の相互作用によりトレーダーの行動をモデル化する．銘柄 A の株を売り買いする n 粒子があるとする．＋粒子が N_1 個と－粒子が N_2 個存在し $n=N_1+N_2$ であるとする．各粒子をトレーダーと考える．＋粒子は株を買いたいと思っているトレーダーとする－粒子は株を売りたいと思っているトレーダー．各ステップはつぎの 1,2,3，からなるとする．

1. 突然変異：n 粒子のなかからランダムに選ばれた 1 粒子の正負が確率 r で反対になる．確率 $1-r$，でもとのままにとどまる ($0\le r\le 1$)．
2. 多数決ルール：3 粒子をランダムに選ぶ．2 粒子が＋(－)，1 粒子が－(＋)，であれば，－(＋) の 1 粒子は＋(－) に変化し価格 S は $1(-1)$ 増える，3 粒子とも＋(－) の場合，粒子の＋－に変化はおきずに価格 S は $3(-3)$ 増加する．
3. フィードバック：S が正（負）の場合，N_1 は確率 $\lambda|S/n|<1$ で $-1(+1)$ 増加する．

このモデルは 2 節でのべたモデルと違い粒子数が負になってしまう可能性があるが，そのようなことはほとんど起きない自然な数値的な設定を考えることができる．時刻 u において各粒子数は $0<N_1(u)$, $N_2(u)$, であり $\lambda|S/n|<1$ であるとすると時刻 $u+1$ での期待値は

$$E\left[\frac{N_1(u+1)-N_1(u)}{\tau n}\right]$$
$$=r\left\{-\frac{N_1(u)}{n}+\frac{N_2(u)}{n}\right\}$$
$$+3\frac{N_1(u)(N_1(u)-1)N_2(u)}{n(n-1)(n-2)}$$
$$-3\frac{N_1(u)N_2(u)(N_2(u)-1)}{n(n-1)(n-2)}-\lambda\frac{S(u)}{n} \tag{4}$$

$$E\left[\frac{S(u+1)-S(u)}{\tau n}\right]$$
$$=3\frac{N_1(u)(N_1(u)-1)(N_1(u)-2)}{n(n-1)(n-2)}$$
$$+3\frac{N_1(u)(N_1(u)-1)N_2(u)}{n(n-1)(n-2)}$$
$$-3\frac{N_2(u)(N_2(u)-1)N_1(u)}{n(n-1)(n-2)}$$
$$-\frac{N_2(u)(N_2(u)-1)(N_2(u)-2)}{n(n-1)(n-2)} \tag{5}$$

となる[15]．フィードバックがなければ ($\lambda=0$)，4 体の相互作用を考えた場合の式(3)において $s=-1/2$ の場合である．n を十分大とし，$\dfrac{N_1(t)-N/2}{n}$ を x_t, $S(t)/n$ を y_t，とし時間スケールを適切にとれば，

$$\frac{d}{dt}x_t=-2rx_t+6x_t\left(\frac{1}{2}+x_t\right)\left(\frac{1}{2}-x_t\right)-\lambda y_t \tag{6}$$
$$=-2\left(r-\frac{3}{4}\right)x_t-6x_t^3-\lambda y_t,$$
$$\frac{d}{dt}y_t=6x_t. \tag{7}$$

により近似される．これは van der Pol 方程式であり $r<3/4$ の場合は極限周期軌道がある．振動論のことばでは発振現象であり[14]，株価は振動するということになる．トレーダーのポジションを考慮に入れると市場での価格のうごきをさらによく理解することができる[18]．現実には確率微分方程式を考えることが自然であり，van der Pol 方程式を時系列モデル化し，データを解析する方法がある[12]．

参考文献

[1] Courtault, J. M. et al.: Louis Bachelier on the centenary of theorie de la speculation, Mathematical

Finance, 10 (2000), 341-353.

[2] Bachelier, L.: Theorie de la speculation, Annales scientifique de l'Ecole Normale Sperieure, 3^{e} serie. 17 (1900), 21-86. (also in Edtions Jacques Gabay).

[3] Hamada, Y.: A method for solving a stochastic differential equation, Progress of Theoretical Physics, 64 (1980), 1127-1137.

[4] 藤原明：日本語はどこから来たか，講談社現代新書 (1981).

[5] Itoh, Y.: On a ruin problem with interaction, Annals of the Institute of Statistical Mathematics, 25 (1973), 635-641.

[6] Itoh, Y.: Random collision models in oriented graphs, Journal of Applied Probability, 16 (1979), 36-44.

[7] Itoh, Y.: Random collision model for random genetic drift and stochastic difference equation, Annals of the Institute of Statistical Mathematics, 36 B (1984), 353-362.

[8] Itoh, Y., Mallows, C. and Shepp, L.: Explicit sufficient invariants for an interacting particle system, Journal of Applied Probability, Vol. 35, No. 3 (1998), 633-641.

[9] Itoh, Y. and Ueda, S.: The Ising model for word ordering rules in natural languages, Physica D, 198 (2005), 333-339.

[10] Maruyama, T. and Nei, M.: Genetic variability maintained by mutation and overdominant selection in finite populations, Genetics, 98 (1982), 441-459.

[11] 向井輝美：集団遺伝学，講談社 (1978).

[12] Ozaki, T.: A local linearization approach to non-linear filtering, International Journal on Control, 57, 1 (1993), 75-96.

[13] Suzuki, M. and Kubo, R.: Dynamics of the Ising model near the critical point I, Journal of the Physical Society of Japan, 24 (1968), 51-60.

[14] 高橋陽一郎：微分方程式入門，東京大学出版会 (1988).

[15] Takahashi, H. and Itoh, Y.: Majority orienting model for the oscillation of market price, European Physical Journal B, 37 (2004), 271-274.

[16] 角田太作：世界の言語と日本語，くろしお出版 (1991).

[17] Tsunoda, T., Ueda, S. and Itoh, Y.: Adpositions in word-order typology, Linguistics, 33 (1994), 741-761.

[18] Yamashita, T. and Itoh, Y.: The oscillation of stock price by majority orienting traders with investment position, Physica A 374 (2007), 764-772.